

[IEEE HOME](#) | [SEARCH IEEE](#) | [SHOP](#) | [WEB ACCOUNT](#) | [CONTACT IEEE](#)[Membership](#) [Publications/Services](#) [Standards](#) [Conferences](#) [Careers/Jobs](#)**IEEE Xplore®**  
RELEASE 1.5Welcome  
United States Patent and Trademark Office[Help](#) [FAQ](#) [Terms](#) [IEEE Peer](#) [Quick Links](#)

» Search

## Welcome to IEEE Xplore®

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

## Tables of Contents

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

## Search

- ☐ By Author
- ☐ Basic
- ☐ Advanced

## Member Services

- ☐ Join IEEE
- ☐ Establish IEEE Web Account
- ☐ Access the IEEE Member Digital Library

 Print FormatYour search matched **11** of **981247** documents.

A maximum of **11** results are displayed, **25** to a page, sorted by **Relevance** in **descending** order.  
You may refine your search by editing the current search expression or entering a new one the text box.  
Then click **Search Again**.

lempe-ziv and search

## Results:

Journal or Magazine = **JNL** Conference = **CNF** Standard = **STD****1 Generalized Lempel-Ziv parsing scheme and its preliminary analysis and average profile***Louchard, G.; Szpankowski, W.;*Data Compression Conference, 1995. DCC '95. Proceedings, 28-30 March 1995  
Page(s): 262 -271[\[Abstract\]](#) [\[PDF Full-Text \(424 KB\)\]](#) **IEEE CNF****2 Asymptotic behavior of the Lempel-Ziv parsing scheme and digital tre***Jacquet, P.; Szpankowski, W.;*Information Theory, 1995. Proceedings., 1995 IEEE International Symposium on  
17-22 Sept. 1995  
Page(s): 14[\[Abstract\]](#) [\[PDF Full-Text \(96 KB\)\]](#) **IEEE CNF****3 On the average redundancy rate of the Lempel-Ziv code with k-error protocol***Reznik, Y.A.; Szpankowski, W.;*Data Compression Conference, 2000. Proceedings. DCC 2000, 28-30 March 2000  
Page(s): 373 -382[\[Abstract\]](#) [\[PDF Full-Text \(100 KB\)\]](#) **IEEE CNF****4 Data compressor decompressor IC***Shah, I.A.; Johnson, B.C.;*Circuits and Systems, 1990., IEEE International Symposium on, 1-3 May 1990  
Page(s): 41 -44 vol.1

[\[Abstract\]](#) [\[PDF Full-Text \(276 KB\)\]](#) **IEEE CNF**

---

**5 Average profile and limiting distribution for a phrase size in the Lempel-Ziv parsing algorithm**

*Louchard, G.; Szpankowski, W.;*

Information Theory, IEEE Transactions on , Volume: 41 Issue: 2 , March 1995

Page(s): 478 -488

[\[Abstract\]](#) [\[PDF Full-Text \(836 KB\)\]](#) **IEEE JNL**

---

**6 Fixed-slope universal lossy data compression**

*En-hui Yang; Zhen Zhang; Berger, T.;*

Information Theory, IEEE Transactions on , Volume: 43 Issue: 5 , Sept. 1997

Page(s): 1465 -1476

[\[Abstract\]](#) [\[PDF Full-Text \(628 KB\)\]](#) **IEEE JNL**

---

**7 A pattern matching approach to image compression**

*Atallah, M.J.; Szpankowski, W.; Genin, Y.;*

Image Processing, 1996. Proceedings., International Conference on , Volume: : 16-19 Sept. 1996

Page(s): 349 -352 vol.2

[\[Abstract\]](#) [\[PDF Full-Text \(668 KB\)\]](#) **IEEE CNF**

---

**8 Pattern matching image compression**

*Atallah, M.; Genin, Y.; Szpankowski, W.;*

Data Compression Conference, 1996. DCC '96. Proceedings , 31 March-3 April

Page(s): 421

[\[Abstract\]](#) [\[PDF Full-Text \(60 KB\)\]](#) **IEEE CNF**

---

**9 Pattern matching image compression: algorithmic and empirical results**

*Atallah, M.; Genin, Y.; Szpankowski, W.;*

Pattern Analysis and Machine Intelligence, IEEE Transactions on , Volume: 21 Issue: 5 , July 1999

Page(s): 614 -627

[\[Abstract\]](#) [\[PDF Full-Text \(1168 KB\)\]](#) **IEEE JNL**

---

**10 Average profile of the Lempel-Ziv scheme for a Markovian source**

*Jing Tang; Szpankowski, W.;*

Information Theory. 1997. Proceedings., 1997 IEEE International Symposium on

June-4 July 1997

Page(s): 311

[\[Abstract\]](#) [\[PDF Full-Text \(92 KB\)\]](#) **IEEE CNF**

---

**11 Software compression in the client/server environment**

*Factor, M.; Sheinwald, D.; Yassour, B.-A.;*

Data Compression Conference, 2001. Proceedings. DCC 2001. , 27-29 March 2

Page(s): 233 -242

[\[Abstract\]](#) [\[PDF Full-Text \(504 KB\)\]](#) **IEEE CNF**

---

[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) | [Advanced Search](#)  
[Join IEEE](#) | [Web Account](#) | [New this week](#) | [OPAC Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Email Alerting](#)  
[No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to Top](#)

Copyright © 2003 IEEE — All rights reserved

[IEEE HOME](#) | [SEARCH IEEE](#) | [SHOP](#) | [WEB ACCOUNT](#) | [CONTACT IEEE](#)[Membership](#) [Publications/Services](#) [Standards](#) [Conferences](#) [Careers/Jobs](#)**IEEE Xplore®**  
RELEASE 1.5Welcome  
United States Patent and Trademark Of[Help](#) [FAQ](#) [Terms](#) [IEEE Peer](#) [Quick Links](#) [Review](#)

&gt;&gt; A

Welcome to IEEE Xplore® [SEARCH RESULTS](#) [\[PDF Full-Text \(276 KB\)\]](#) [PREVIOUS](#) [NEXT](#) [DOWNLOAD CITATION](#)

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

**Tables of Contents**

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

**Search**

- ☐ By Author
- ☐ Basic
- ☐ Advanced

**Member Services**

- ☐ Join IEEE
- ☐ Establish IEEE Web Account
- ☐ Access the IEEE Member Digital Library

 [Print Format](#)

## Data compressor decompressor IC

Shah, I.A. Johnson, B.C.

North American Philips Corp., Briarcliff Manor, NY ;

*This paper appears in: **Circuits and Systems, 1990., IEEE International Symposium on***

Meeting Date: 05/01/1990 -05/03/1990

Publication Date: 1-3 May 1990

Location: New Orleans, LA , USA

On page(s): 41-44 vol.1

References Cited: 3

INSPEC Accession Number: 3832467

**Abstract:**

The data compressor decompressor (DCD) IC, a VLSI implementation of a version of the Lempel-Ziv (L-Z) compression algorithm, is discussed. The IC is suitable for high-speed, lossless compression of digital data such as text and images. The single-pass lossless compression scheme adapts to the statistics of the data to be processed. The authors outline the L-Z compression method, illustrate a solution to its search requirement, and discuss the chip architecture and features

**Index Terms:**

VLSI data compression picture processing signal processing equipment L-Z compression method Lempel-Ziv compression algorithm VLSI chip architecture data compressor decompressor images lossless compression search requirement single-pass lossless compression scheme text

**Documents that cite this document**

Select link to view other documents in the database that cite this one.

[SEARCH RESULTS](#) [\[PDF Full-Text \(276 KB\)\]](#) [PREVIOUS](#) [NEXT](#) [DOWNLOAD CITATION](#)

[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) | [Advanced Search](#) | [Join IEEE](#) | [Web Account](#) | [New this week](#) | [OPAC Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Errata](#) | [No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to Top](#)



[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent & Trademark Office



Try the *new* Portal design

Give us your opinion after using it.

## Search Results

Search Results for: [(((compress or compact or encode) <near/2> data) <paragraph> (field or column or attribution)) and (decompress or decode) and ((compress or encode or compact) <sentence> search) ]

Found 28 of 122,783 searched.

## Search within Results



[> Advanced Search](#)

[> Search Help/Tips](#)

Sort by: [Title](#) [Publication](#) [Publication Date](#) [Score](#)  [Binder](#)

Results 1 - 20 of 28 [short listing](#)



1 2



### 1 [Profile-guided code compression](#)

100%



Saumya Debray , William Evans

**ACM SIGPLAN Notices , Proceedings of the ACM SIGPLAN 2002 Conference on Programming language design and implementation** May 2002

Volume 37 Issue 5

As computers are increasingly used in contexts where the amount of available memory is limited, it becomes important to devise techniques that reduce the memory footprint of application programs while leaving them in an executable form. This paper describes an approach to applying data compression techniques to reduce the size of infrequently executed portions of a program. The compressed code is decompressed dynamically (via software) if needed, prior to execution. The use of data compression t ...

### 2 [Query optimization in compressed database systems](#)

100%








Zhiyuan Chen , Johannes Gehrke , Flip Korn

**ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on Management of data** May 2001

Volume 30 Issue 2

Over the last decades, improvements in CPU speed have outpaced improvements in main memory and disk access rates by orders of magnitude, enabling the use of data compression techniques to improve the performance of database systems. Previous work describes the benefits of compression for numerical attributes, where data is stored in compressed format on disk. Despite the abundance of string-valued attributes in relational schemas there is little

work on compression for string attributes in a ...

- 3 Tool support for architectural decisions in embedded systems: CoCo: a hardware/software platform for rapid prototyping of code compression technologies 100%  
 Haris Lekatsas , Jörg Henkel , Srimat Chakradhar , Venkata Jakkula , Murugan Sankaradass  
**Proceedings of the 40th conference on Design automation** June 2003  
In recent years instruction code compression/decompression technologies have emerged as an efficient way to a) reduce the memory usage of an embedded system, b) to improve performance through effectively higher bandwidths and/or to c) reduce the overall power consumption of a system processing compressed code. We have presented efficient code compression/decompression techniques and architectures in the past. For the commercialization phase, we designed a novel hardware/software code compression ...
- 4 Parallel algorithms for data compression 100%  
 M. E. Gonzalez Smith , J. A. Storer  
**Journal of the ACM (JACM)** April 1985  
Volume 32 Issue 2  
Parallel algorithms for data compression by textual substitution that are suitable for VLSI implementation are studied. Both &ldquo;static&rdquo; and &ldquo;dynamic&rdquo; dictionary schemes are considered.
- 5 SAMC - efficient semi-adaptive data compression 100%  
 Edward Hatton  
**Proceedings of the 1995 conference of the Centre for Advanced Studies on Collaborative research** November 1995  
Universal noiseless coding is of considerable interest to industry for the purposes of data reduction in order to store or transmit large volumes of typically textual data. Compression schemes have evolved from simple memoryless Huffman coding, to the Lempel-Ziv family of dictionary compression, to the current Markov or statistical modelling. This evolution has resulted in successively better compression, at an increased cost of execution time and RAM requirements. Bell, Cleary, and Moffat's Mar ...
- 6 Data compression 99%  
 Debra A. Lelewer , Daniel S. Hirschberg  
**ACM Computing Surveys (CSUR)** September 1987  
Volume 19 Issue 3  
This paper surveys a variety of data compression methods spanning almost 40 years of research, from the work of Shannon, Fano, and Huffman in the late 1940s to a technique developed in 1986. The aim of data compression is to reduce redundancy in stored or communicated data, thus increasing effective data density. Data compression has important application in the areas of file storage and distributed systems. Concepts from information theory as they relate to the goals and evaluation of data ...
- 7 Dictionary-based order-preserving string compression 99%  
 Gennady Antoshenkov  
**The VLDB Journal &mdash; The International Journal on Very Large Data Bases**  
February 1997

## Volume 6 Issue 1

As no database exists without indexes, no index implementation exists without order-preserving key compression, in particular, without prefix and tail compression. However, despite the great potentials of making indexes smaller and faster, application of general compression methods to ordered data sets has advanced very little. This paper demonstrates that the fast dictionary-based methods can be applied to order-preserving compression almost with the same freedom as in the general case. The pro ...

8 Modeling for text compression

98%



Timothy Bell , Ian H. Witten , John G. Cleary

**ACM Computing Surveys (CSUR)** December 1989

Volume 21 Issue 4

The best schemes for text compression use large models to help them predict which characters will come next. The actual next characters are coded with respect to the prediction, resulting in compression of information. Models are best formed adaptively, based on the text seen so far. This paper surveys successful strategies for adaptive modeling that are suitable for use in practical text compression systems. The strategies fall into three main classes: finite-context modeling, i ...

9 XML indexing and compression: XPRESS: a queriable compression for XML data

97%



Jun-Ki Min , Myung-Jae Park , Chin-Wan Chung

**Proceedings of the 2003 ACM SIGMOD international conference on on Management of data** June 2003

Like HTML, many XML documents are resident on native file systems. Since XML data is irregular and verbose, the disk space and the network bandwidth are wasted. To overcome the verbosity problem, the research on compressors for XML data has been conducted. However, some XML compressors do not support querying compressed data, while other XML compressors which support querying compressed data blindly encode tags and data values using predefined encoding methods. Thus, the query performance on com ...

10 The performance advantage of applying compression to the memory system

96%



Nihar R. Mahapatra , Jiangjiang Liu , Krishnan Sundaresan

**ACM SIGPLAN Notices , Proceedings of the workshop on Memory system performance** June 2002

Volume 38 Issue 2 supplement

The memory system stores information comprising primarily instructions and data and secondarily address information, such as cache tag fields. It interacts with the processor by supporting related traffic (again comprising addresses, instructions, and data). Continuing exponential growth in processor performance, combined with technology, architecture, and application trends, place enormous demands on the memory system to permit this information storage and exchange at a high-enough performance ...

11 Hardware-Assisted Data Compression for Energy Minimization in Systems with Embedded

95%

Processors

L. Benini , D. Bruni , A. Macii , E. Macii

**Proceedings of the conference on Design, automation and test in Europe** March 2002

In this paper, we suggest hardware-assisted data compression as a tool for reducing energy consumption of core-based embedded systems. We propose a novel and efficient

architecture for on-the-fly data compression and decompression whose field of operation is the cache-to-memory path. Uncompressed cache lines are compressed before they are written back to main memory, and decompressed when cache refills take place. We explore two classes of compression methods, profile-driven and differential, since ...

## 12 Optimal VLSI circuits for sorting

95%



Richard Cole , Alan Siegel

**Journal of the ACM (JACM)** October 1988

Volume 35 Issue 4

This work describes a large number of constructions for sorting  $N$  integers in the range  $[0, M - 1]$ , for  $N \leq M \leq N^2$ , for the standard VLSI bit model. Among other results, we attain: VLSI sorter constructions that are within a constant factor of optimal size, for all  $M$  and almost all running times  $T$ . a ...

## 13 A very fast algorithm for RAM compression

93%



Luigi Rizzo

**ACM SIGOPS Operating Systems Review** April 1997

Volume 31 Issue 2

Compressed virtual memory systems have been suggested, and in some cases implemented, to improve the effectiveness of use of physical RAM. However, most proposals and/or implementations are based on adaptive compression algorithms which achieve good compression ratios, but are slow compared to a local disk. Hence, they can only give some advantage with very slow (e.g. network-mounted) swap devices. In this paper we show that in many cases memory pages contain highly compressible data, with a ver ...

## 14 LeZi-update: an information-theoretic framework for personal mobility tracking in PCS networks

87%



Amiya Bhattacharya , Sajal K. Das

**Wireless Networks** March 2002

Volume 8 Issue 2/3

The complexity of the mobility tracking problem in a cellular environment has been characterized under an information-theoretic framework. Shannon's entropy measure is identified as a basis for comparing user mobility models. By building and maintaining a dictionary of individual user's *path* updates (as opposed to the widely used location updates), the proposed adaptive on-line algorithm can learn subscribers' profiles. This technique evolves out of the concepts of lossless compression. T ...

## 15 A scheme for data compression in supercomputers

87%








M. A. Bassiouni , N. Ranganathan , A. Mukherjee

**Proceedings of the 1988 ACM/IEEE conference on Supercomputing** November 1988

There is a growing recognition of the importance of efficient coding and data compression schemes in supercomputing centers and in networks of high-speed computing machines. Recently, there has been a considerable interest in arithmetic coding as a promising technique for reducing the cost of data storage and transmission. In this paper, we present a compression algorithm that is tailored to utilize the enormous speed and memory size of supercomputers and which utilizes an enhanced ...



- 16 LeZi-update: an information-theoretic approach to track mobile users in PCS networks 87%  
 Amiya Bhattacharya , Sajal K. Das  
**Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking** August 1999
- 17 Query evaluation techniques for large databases 83%  
 Goetz Graefe  
**ACM Computing Surveys (CSUR)** June 1993  
Volume 25 Issue 2  
Database management systems will continue to manage large data volumes. Thus, efficient algorithms for accessing and manipulating large sets and sequences will be required to provide acceptable performance. The advent of object-oriented and extensible database systems will not solve this problem. On the contrary, modern data models exacerbate the problem: In order to manipulate large sets of complex objects as efficiently as today's database systems manipulate simple records, query-processi ...
- 18 Graded codes by G-sets 82%  
 J. Peralta , B. Torrecillas  
**ACM SIGSAM Bulletin** September 1999  
Volume 33 Issue 3
- 19 Evolvable hardware chips for industrial applications 82%  
 Tetsuya Higuchi , Nobuki Kajihara  
**Communications of the ACM** April 1999  
Volume 42 Issue 4
- 20 Optimizing methods: Data compression techniques for economic processing of large commercial files 82%  
 James E. Mulford , Richard K. Ridall  
**Proceedings of the 1971 international ACM SIGIR conference on Information storage and retrieval** April 1971  
The application of compact coding, differencing and other techniques to indexed sequential files is discussed. The effects on system performance are discussed and reductions of almost 80% in mass storage requirements for a particular file are reported.

---

Results 1 - 20 of 28    short listing



[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent & Trademark Office



Try the **new** Portal design

Give us your opinion after using it.

## Search Results

Search Results for: [(((compress or compact or encode) <near/2> data) <paragraph> (field or column or attribution)) and (decompress or decode) and ((compress or encode or compact) <sentence> search) ]


Found 28 of 122,783 searched.

## Search within Results



[> Advanced Search](#)

[> Search Help/Tips](#)

Sort by: [Title](#) [Publication](#) [Publication Date](#) [Score](#)  [Binder](#)

Results 21 - 28 of 28 [short listing](#)



[1](#) [2](#)



### 21 [Techniques for FPGA implementation of video compression systems](#)

82%



Brian Schoner , John Villasenor , Steve Molloy , Rajeev Jain

**Proceedings of the 1995 ACM third international symposium on Field-programmable gate arrays** February 1995

Real-time video compression is a challenging subject for FPGA implementation because it typically has a large computational complexity and requires high data throughput. Previous implementations have used parallel banks of FPGAs or DSPs to meet these requirements. Using design techniques that maximize FPGA utilization, we have implemented two video compression systems, each of which uses a single FPGA. In this first system, algorithmic optimizations are made to create a low-complexity imple ...

### 22 [Demonstration session 1: Clear face analysis from MPEG compressed video](#)

80%









Ling-Yu Duan , Qi Tian

**Proceedings of the tenth ACM international conference on Multimedia** December 2002

In this demonstration, we present a system to analyze the clear degree of faces present in MPEG compressed video of Head-and-Shoulders style. The proposed system consists of three hierarchical modules: low-level features extraction, robust face tracking, and clear faces selection. We have integrated the core algorithm into an Automated Transaction Service (ATS) surveillance system. The Incremental Focus of Attention (IFA) architecture is taken to combine pixel domain processing with compressed d ...

### 23 [Session P11: visualization systems and image-based visualization: Sea of images](#)

80%

-  Daniel G. Aliaga , Thomas Funkhouser , Dimah Yanovsky , Ingrid Carlbom  
**Proceedings of the conference on Visualization '02** October 2002  
A long-standing research problem in computer graphics is to reproduce the visual experience of walking through a large photorealistic environment interactively. On one hand, traditional geometry-based rendering systems fall short of simulating the visual realism of a complex environment. On the other hand, image-based rendering systems have to date been unable to capture and store a sampled representation of a large environment with complex lighting and visibility effects. In this paper, we prese ...
- 24 A survey on wavelet applications in data mining 77%  
 Tao Li , Qi Li , Shenghuo Zhu , Mitsunori Ogihara  
**ACM SIGKDD Explorations Newsletter** December 2002  
Volume 4 Issue 2  
Recently there has been significant development in the use of wavelet methods in various data mining processes. However, there has been written no comprehensive survey available on the topic. The goal of this is paper to fill the void. First, the paper presents a high-level data-mining framework that reduces the overall process into smaller components. Then applications of wavelets for each component are reviewd. The paper concludes by discussing the impact of wavelets on data mining research an ...
- 25 Sequentially encoded data structures that support bidirectional scanning 77%  
 Robert J. Lechner  
**ACM SIGARCH Computer Architecture News , Proceedings of the 2nd annual symposium on Computer architecture** December 1974  
Volume 3 Issue 4
- 26 Access methods for text 77%  
 Chris Faloutsos  
**ACM Computing Surveys (CSUR)** March 1985  
Volume 17 Issue 1  
This paper compares text retrieval methods intended for office systems. The operational requirements of the office environment are discussed, and retrieval methods from database systems and from information retrieval systems are examined. We classify these methods and examine the most interesting representatives of each class. Attempts to speed up retrieval with special purpose hardware are also presented, and issues such as approximate string matching and compression are discussed. A quali ...
- 27 Inplace run-length 2d compressed search 77%  
 Amihood Amir , Gad M. Landau , Dina Sokol  
**Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms**  
February 2000
- 28 Software/hardware co-design implementation for fractal image compression 77%  
 Tai-Chi Lee , Patrick Robinson , Michael Gubody , Erik Henne  
**Proceedings of the 37th annual Southeast regional conference (CD-ROM)** April 1999
-

Results 21 - 28 of 28   [short listing](#)



1 2



---

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2003 ACM, Inc.

[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent &amp; Trademark Office

Try the *new* Portal design

Give us your opinion after using it.


## Search Results

Search Results for: [(((compress or compact or encode) <near/2> data) <paragraph> (field or column or attribution)) and (decompress or decode) and ((compress or encode or compact) <sentence> search) ]

Found 28 of 122,783 searched.

Search within Results

[> Advanced Search](#)[> Search Help/Tips](#)









Sort by: [Title](#) [Publication](#) [Publication Date](#) [Score](#)  [Binder](#)

Results 21 - 28 of 28 [long listing](#)

Prev  
Page

1 2

Next  
Page

- 21 [Techniques for FPGA implementation of video compression systems](#) 82%  
 Brian Schoner , John Villasenor , Steve Molloy , Rajeev Jain  
**Proceedings of the 1995 ACM third international symposium on Field-programmable gate arrays** February 1995
- 22 [Demonstration session 1: Clear face analysis from MPEG compressed video](#) 80%  
 Ling-Yu Duan , Qi Tian  
**Proceedings of the tenth ACM international conference on Multimedia** December 2002
- 23 [Session P11: visualization systems and image-based visualization: Sea of images](#) 80%  
 Daniel G. Aliaga , Thomas Funkhouser , Dimah Yanovsky , Ingrid Carlbom  
**Proceedings of the conference on Visualization '02** October 2002
- 24 [A survey on wavelet applications in data mining](#) 77%  
 Tao Li , Qi Li , Shenghuo Zhu , Mitsunori Ogihara  
**ACM SIGKDD Explorations Newsletter** December 2002  
Volume 4 Issue 2
- 25 [Sequentially encoded data structures that support bidirectional scanning](#) 77%  
 Robert J. Lechner  
**ACM SIGARCH Computer Architecture News , Proceedings of the 2nd annual symposium on Computer architecture** December 1974  
Volume 3 Issue 4
- 26 [Access methods for text](#) 77%  
 Chris Faloutsos  
**ACM Computing Surveys (CSUR)** March 1985  
Volume 17 Issue 1
- 27 [Inplace run-length 2d compressed search](#) 77%  
 Amihood Amir , Gad M. Landau , Dina Sokol  
**Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms**  
February 2000
- 28 [Software/hardware co-design implementation for fractal image compression](#) 77%  
 Tai-Chi Lee , Patrick Robinson , Michael Gubody , Erik Henne  
**Proceedings of the 37th annual Southeast regional conference (CD-ROM)** April 1999

---

Results 21 - 28 of 28      [long listing](#)



1 2



. Inc.



[> home](#) [> about](#) [> feedback](#) [> login](#)

US Patent & Trademark Office



Try the *new* Portal design

Give us your opinion after using it.

## Search Results

Search Results for: [(((compress or compact or encode) <near/2> data) <paragraph> (field or column or attribution)) and (decompress or decode) and ((compress or encode or compact) <sentence> search) ]

Found 28 of 122,783 searched.

## Search within Results



[> Advanced Search](#)

[> Search Help/Tips](#)

Sort by: [Title](#) [Publication](#) [Publication Date](#) [Score](#) [Binder](#)

Results 1 - 20 of 28 [long listing](#)

Prev Page 1 2 Next Page

- 1 [Profile-guided code compression](#) 100%

Saumya Debray , William Evans

**ACM SIGPLAN Notices , Proceedings of the ACM SIGPLAN 2002 Conference on Programming language design and implementation May 2002**

Volume 37 Issue 5
- 2 [Query optimization in compressed database systems](#) 100%

Zhiyuan.Chen , Johannes Gehrke , Flip Korn

**ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on Management of data May 2001**

Volume 30 Issue 2
- 3 [Tool support for architectural decisions in embedded systems: CoCo: a hardware/software platform for rapid prototyping of code compression technologies](#) 100%

Haris Lekatsas , Jörg Henkel , Srimat Chakradhar , Venkata Jakkula , Murugan Sankaradass









**Proceedings of the 40th conference on Design automation June 2003**
- 4 [Parallel algorithms for data compression](#) 100%









M. E. Gonzalez Smith , J. A. Storer

**Journal of the ACM (JACM) April 1985**

Volume 32 Issue 2



- 5 SAMC - efficient semi-adaptive data compression 100%  
 Edward Hatton  
**Proceedings of the 1995 conference of the Centre for Advanced Studies on Collaborative research** November 1995
  
- 6 Data compression 99%  
 Debra A. Lelewer , Daniel S. Hirschberg  
**ACM Computing Surveys (CSUR)** September 1987  
Volume 19 Issue 3
  
- 7 Dictionary-based order-preserving string compression 99%  
 Gennady Antoshenkov  
**The VLDB Journal &mdash; The International Journal on Very Large Data Bases**  
February 1997  
Volume 6 Issue 1
  
- 8 Modeling for text compression 98%  
 Timothy Bell , Ian H. Witten , John G. Cleary  
**ACM Computing Surveys (CSUR)** December 1989  
Volume 21 Issue 4
  
- 9 XML indexing and compression: XPRESS: a queriable compression for XML data 97%  
 Jun-Ki Min , Myung-Jae Park , Chin-Wan Chung  
**Proceedings of the 2003 ACM SIGMOD international conference on on Management of data** June 2003
  
- 10 The performance advantage of applying compression to the memory system 96%  
 Nihar R. Mahapatra , Jiangjiang Liu , Krishnan Sundaresan  
**ACM SIGPLAN Notices , Proceedings of the workshop on Memory system performance** June 2002  
Volume 38 Issue 2 supplement
  
- 11 Hardware-Assisted Data Compression for Energy Minimization in Systems with Embedded Processors 95%  
 L. Benini , D. Bruni , A. Macii , E. Macii  
**Proceedings of the conference on Design, automation and test in Europe** March 2002
  
- 12 Optimal VLSI circuits for sorting 95%  
 Richard Cole , Alan Siegel  
**Journal of the ACM (JACM)** October 1988  
Volume 35 Issue 4
  
- 13 A very fast algorithm for RAM compression 93%

-  Luigi Rizzo  
**ACM SIGOPS Operating Systems Review** April 1997  
Volume 31 Issue 2
- 14 LeZi-update: an information-theoretic framework for personal mobility tracking in PCS networks 87%  
 Amiya Bhattacharya , Sajal K. Das  
**Wireless Networks** March 2002  
Volume 8 Issue 2/3
- 15 A scheme for data compression in supercomputers 87%  
 M. A. Bassiouni , N. Ranganathan , A. Mukherjee  
**Proceedings of the 1988 ACM/IEEE conference on Supercomputing** November 1988
- 16 LeZi-update: an information-theoretic approach to track mobile users in PCS networks 87%  
 Amiya Bhattacharya , Sajal K. Das  
**Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking** August 1999
- 17 Query evaluation techniques for large databases 83%  
 Goetz Graefe  
**ACM Computing Surveys (CSUR)** June 1993  
Volume 25 Issue 2
- 18 Graded codes by G-sets 82%  
 J. Peralta , B. Torrecillas  
**ACM SIGSAM Bulletin** September 1999  
Volume 33 Issue 3
- 19 Evolvable hardware chips for industrial applications 82%  
 Tetsuya Higuchi , Nobuki Kajihara  
**Communications of the ACM** April 1999  
Volume 42 Issue 4
- 20 Optimizing methods: Data compression techniques for economic processing of large commercial files 82%  
 James E. Mulford , Richard K. Ridall  
**Proceedings of the 1971 international ACM SIGIR conference on Information storage and retrieval** April 1971

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2003 ACM, Inc.

[IEEE HOME](#) | [SEARCH IEEE](#) | [SHOP](#) | [WEB ACCOUNT](#) | [CONTACT IEEE](#)[Membership](#) [Publications/Services](#) [Standards](#) [Conferences](#) [Careers/Jobs](#)**IEEE Xplore®**  
RELEASE 1.5Welcome  
United States Patent and Trademark Of[Help](#) [FAQ](#) [Terms](#) [IEEE Peer](#) [Quick Links](#) [» Se](#)  
[Review](#)

## Welcome to IEEE Xplore®

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out


## Tables of Contents

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

## Search

- ☐ By Author
- ☐ Basic
- ☐ Advanced

## Member Services

- ☐ Join IEEE
- ☐ Establish IEEE Web Account
- ☐ Access the IEEE Member Digital Library
-  Print Format

Your search matched **4** of **983096** documents.

A maximum of **4** results are displayed, **25** to a page, sorted by **Relevance** in **descending** order.  
You may refine your search by editing the current search expression or entering a new one the text b  
Then click **Search Again**.

((compress or encode or compact) &lt;paragraph&gt; search) and ((decompress or decode) &lt;p

[Search Again](#)

## Results:

Journal or Magazine = **JNL** Conference = **CNF** Standard = **STD****1 A high speed digital data separator design using real time DSP for dis applications***Kim, B.; Greco, J.D.; Helman, D.N.; Ngo, H.; Yang, H.C.; Wu, W.-C.S.; Chowd, R.F.;*

Circuits and Systems, 1992. ISCAS '92. Proceedings., 1992 IEEE International Symposium on , Volume: 2 , 3-6 May 1992

Page(s): 625 -628 vol.2

[\[Abstract\]](#) [\[PDF Full-Text \(308 KB\)\]](#) **IEEE CNF****2 Order preserving string compression***Antoshenkov, G.; Lomet, D.; Murray, J.;*

Data Engineering, 1996. Proceedings of the Twelfth International Conference o Feb.-1 March 1996

Page(s): 655 -663

[\[Abstract\]](#) [\[PDF Full-Text \(804 KB\)\]](#) **IEEE CNF****3 Practical constructions of L-restricted alphabetic prefix codes***Laber, E.S.; Milidiu, R.L.; Pessoa, A.A.;*

String Processing and Information Retrieval Symposium, 1999 and Internation Workshop on Groupware , 22-24 Sept. 1999

Page(s): 115 -119

[\[Abstract\]](#) [\[PDF Full-Text \(92 KB\)\]](#) **IEEE CNF****4 Low-complexity global motion estimation from P-frame motion vecto MPEG-7 applications**

*Smolic, A.; Hoeynck, M.; Ohm, J.-R.;*

Image Processing, 2000. Proceedings. 2000 International Conference on , Volu  
10-13 Sept. 2000

Page(s): 271 -274 vol.2

[\[Abstract\]](#) [\[PDF Full-Text \(464 KB\)\]](#) **IEEE CNF**

---

[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) | [Advanced Search](#)  
[Join IEEE](#) | [Web Account](#) | [New this week](#) | [OPAC Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Email Alerting](#)  
[No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to Top](#)

Copyright © 2003 IEEE — All rights reserved

[IEEE HOME](#) | [SEARCH IEEE](#) | [SHOP](#) | [WEB ACCOUNT](#) | [CONTACT IEEE](#)[Membership](#) | [Publications/Services](#) | [Standards](#) | [Conferences](#) | [Careers/Jobs](#)**IEEE Xplore®**  
RELEASE 1.5Welcome  
United States Patent and Trademark Office[Help](#) | [FAQ](#) | [Terms](#) | [IEEE Peer Review](#) | [Quick Links](#) [Welcome to IEEE Xplore®](#) | [SEARCH RESULTS](#) | [\[PDF Full-Text \(804 KB\)\]](#) | [PREVIOUS](#) | [NEXT](#) | [DOWNLOAD CITATION](#)

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

**Tables of Contents**

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

**Search**

- ☐ By Author
- ☐ Basic
- ☐ Advanced

**Member Services**

- ☐ Join IEEE
- ☐ Establish IEEE Web Account
- ☐ Access the IEEE Member Digital Library

 [Print Format](#)

---

## Order preserving string compression

Antoshenkov, G. Lomet, D. Murray, J.

Digital Equipment Corp., Maynard, MA;

*This paper appears in: Data Engineering, 1996. Proceedings of the Twelfth International Conference on*

Meeting Date: 02/26/1996 -03/01/1996

Publication Date: 26 Feb-1 Mar 1996

Location: New Orleans, LA , USA

On page(s): 655-663

References Cited: 11

INSPEC Accession Number: 5242668

---

### Abstract:

Order-preserving compression can improve sorting and searching performance and hence the performance of database systems. We describe a new parsing (tokenization) technique that can be applied to variable-length "keys", producing substantial compression. It can both compress and decompress data, permitting variable lengths for dictionary entries and compressed forms. The key notion is to partition the space of strings into ranges encoding the common prefix of each range. We illustrate our method with padding character compression for multi-field keys, demonstrating the dramatic gains possible. A specific version of the method has been implemented in Digital's Ingres relational database system to enable effective multi-field compression

---

### Index Terms:

data compression encoding relational databases sorting Digital Rdb relational database system compressed forms data decompression database systems performance multi-field keys order-preserving string compression padding character compression parsing technique range common prefix encoding searching performance string-space partitioning tokenization technique variable-length dictionary entries variable-length keys

---

### Documents that cite this document

Select link to view other documents in the database that cite this one.

---

---

[SEARCH RESULTS](#) | [\[PDF Full-Text \(804 KB\)\]](#) | [PREVIOUS](#) | [NEXT](#) | [DOWNLOAD CITATION](#)

---



[> home](#) [> about](#) [> feedback](#) [> logout](#)

US Patent &amp; Trademark Office

Citation

## ACM Transactions on Information Systems (TOIS)

[> archive](#)Volume 15 , Issue 2 (April 1997) [> toc](#)

### Making a digital library: the contents of the CORE project

#### Authors

Richard Entlich Cornell Univ., Ithaca, NY  
Jan Olsen Cornell Univ., Ithaca, NY  
Lorrin Garson American Chemical Society  
Michael Lesk Bellcore  
Lorraine Normore Chemical Abstracts Service  
Stuart Weibel OCLC, Dublin, OH


#### Publisher

ACM Press New York, NY, USA

Pages: 103 - 123 Periodical-Issue-Article

Year of Publication: 1997


ISSN:1046-8188

 <http://doi.acm.org/10.1145/248625.248627> (Use this link to Bookmark this page)

[> full text](#) [> abstract](#) [> references](#) [> citings](#) [> index terms](#)  
[> review](#) [> peer to peer](#)


---


[> Discuss](#) [> Similar](#) [> Review this Article](#)

 [Save to Binder](#)

[> BibTex Format](#)

---

 **FULL TEXT:**  [Access Rules](#)

 **pdf 1.50 MB**

 **ABSTRACT**

The CORE (Chemical Online Retrieval Experiment) project is a library of primary journal articles in chemistry. Any library has an inside and an outside; in this article we describe the inside of the library and the methods for building the system and accumulating the database. A later



article will describe the outside (user experiences). Among electronic-library projects, the CORE project is unusual in that it has both ASCII derived from typesetting and image data for all its pages, and among experimental electronic-library projects, it is unusually large. We describe here (a) the processes of scanning and analyzing about 400,000 pages of primary journal material, (b) the conversion of a similar amount of textual database material, (c) the linking of these two data sources, and (d) the indexing of the text material.

## ➤ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

- 1 AAP. 1987. Standard for electronic manuscript preparation and markup. Version 2.0. Assoc. of American Publishers, New York.
- 2 BAIRD, H.S. 1987. The skew angle of printed documents. In Proceedings of the SPSE 40th Conference on Hybrid Imaging Systems. Society of Photographic Scientists and Engineers, Rochester, N.Y., 21-24.
- 3 BIRMINGHAM, W. P., DRABENSTOTT, K. M., FROST, C. O., WARNER, A. J., AND WILLIS, K. 1994. The University of Michigan Digital Library: This is not your father's library. In Digital Libraries '94 Proceedings. Texas A & M Univ., College Station, Tex., 53-60.
- 4 James H. Coombs , Allen H. Renear , Steven J. DeRose, Markup systems and the future of scholarly text processing, Communications of the ACM, v.30 n.11, p.933-947, Nov. 1987
- 5 Dennis E. Egan , Michael E. Lesk , R. Daniel Ketchum , Carol C. Lochbaum , Joel R. Remde , Michael Littman , Thomas K. Landauer, Hypertext for the electronic library?: CORE sample results, Proceedings of the third annual ACM conference on Hypertext, p.299-312, December 15-18, 1991, San Antonio, Texas, United States
- 6 Dennis E. Egan , Joel R. Remde , Thomas K. Landauer , Carol C. Lochbaum , L. M. Gomez, Behavioral evaluation and analysis of a hypertext browser, ACM SIGCHI Bulletin, v.20 n.SI, p.205-210
- 7 FERGUSON, E. 1992. Engineering and the Mind's Eye. MIT Press, Cambridge, Mass.
- 8 FLETCHER, L. A. AND KASTURI, R. 1987. Segmentation of binary images into text strings and graphics. In Proceedings of SPIE Conference on Applications of Artificial Intelligence V. Vol. 786. SPIE, Bellingham, Wash., 533-540.
- 9 Melia M. Hoffman , Lawrence O'Gorman , Guy A. Story , James Q. Arnold , Nina H. Macdonald, The RightPages services: an image-based electronic library, Journal of the American Society for Information Science, v.44 n.8, p.446-452, Sept. 1993
- 10 KIBBEY, M. AND EVANS, N.H. 1989. The network is the library. EDUCOM Rev. 24, 3, 15-20.
- 11 KIRSTEIN, P. AND MONTASSER-KOHSARI, C. 1995. The C-ODA project: Experiences and tools. Comput. J. 38, 8, 670-680.
- 12 KLING, R. AND ELLIOT, M. 1994. Digital library design for organizational usability. In Digital Libraries '94 Proceedings. 146-155.
- 13 LOVE, R.A. 1986. Chemical Journals Online (CJO)--The new full-text database through STN International. In Proceedings Online '86. Online, Inc., Weston, Conn., 149-151.
- 14 LYNCH, C.A. 1989. From telecommunications to networking: The MELVYL online union catalog and the development of intercampus networks at the University of California. Libr. Hi

Tech 7, 2, 61-83.

- 15 MCKNIGHT, C. 1993. Electronic journals--Past, present., and future? ASLIB Proc. 45, 7-10.
- 16 MIKSA, F. L. AND DOTY, P. 1994. Intellectual realities and the digital library. In Digital Libraries '94 Proceedings. 1-5.
- 17 John K. Ousterhout, Tcl and the Tk toolkit, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1994
- 18 SCHATZ, B., BISHOP, A., MISCHO, W., AND HARDIN, J. 1994. Digital library infrastructure for a university engineering community. In Digital Libraries '94 Proceedings. 21-24.
- 19 SRIHARI, S. N., LAM, S. W., HULL, J. J., SRIHARI, R. K., AND GOVINDARAJU, V. 1994. Intelligent data retrieval from raster images of documents. In Digital Libraries '94 Proceedings. 34-40.
- 20 STERN, B.T. 1990. ADONIS--A vision of the future. In Interlending and Document Supply, G. P. Cornish, and A. Gallico, Eds. British Library, 23-33.
- 21 Dacheng Wang , Sargur N. Srihari, Classification of newspaper image blocks using texture analysis, Computer Vision, Graphics, and Image Processing, v.47 n.3, p.327-352, Sept. 1989

#### ↗ CITINGS 2

Gene Golovchinsky , Morgan N. Price , Bill N. Schilit, From reading to retrieval: freeform ink annotations as queries, Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, p.19-25, August 15-19, 1999, Berkeley, California, United States

Bill N. Schilit , Morgan N. Price , Gene Golovchinsky, Digital library information appliances, Proceedings of the third ACM conference on Digital libraries, p.217-226, June 23-26, 1998, Pittsburgh, Pennsylvania, United States

#### ↗ INDEX TERMS

##### Primary Classification:

H. Information Systems

↗ H.3 INFORMATION STORAGE AND RETRIEVAL

↗ H.3.6 Library Automation

↗ Subjects: Large text archives

##### Additional Classification:

H. Information Systems

↗ H.3 INFORMATION STORAGE AND RETRIEVAL

↗ H.5 INFORMATION INTERFACES AND PRESENTATION (I.7)

↗ H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)

↗ Subjects: Interaction styles (e.g., commands, menus, forms, direct manipulation)

I. Computing Methodologies

↗ I.4 IMAGE PROCESSING AND COMPUTER VISION

##### General Terms:

Algorithms, Design, Experimentation, Human Factors

##### Keywords:

image segmentation

#### ↗ REVIEW

"Friedrich Gebhardt"

The goal of CORE is to study the problems of text retrieval with huge amounts of data (many gigabytes) including pictures. The data include several years of 20 journals published by the American Chemical Society (ACS). The text is available in the internal format used by the ACS; the pictures (almost entirely graphs) have to be identified in the printed journals or on films, scanned, and assigned to the proper places within the text. The paper explains the tasks of figure extraction and text acquisition in sufficient detail, giving examples. In addition, the SCEPTER user interface and the Newton search engine are sketched, again with illustrations.  
*Online Computing Reviews Service*

✦ **Peer to Peer - Readers of this Article have also read:**

The knowledge complexity of interactive proof-systems

**Proceedings of the seventeenth annual ACM symposium on Theory of computing**

S Goldwasser , S Micali , C Rackoff

New Products

**Linux Journal** 1996, 27es

CORPORATE Linux Journal Staff

Book Review: IPv6: The New Internet Protocol

**Linux Journal** 1996, 25es

CORPORATE Linux Journal Staff

Programming languages for distributed computing systems

**ACM Computing Surveys (CSUR)** 21, 3

Henri E. Bal , Jennifer G. Steiner , Andrew S. Tanenbaum

Book Review: Civilizing Cyberspace

**Linux Journal** 1996, 28es

CORPORATE Linux Journal Staff

## Welcome to IEEE Xplore™

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

## Tables of Contents

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

## Search

- ☐ By Author
- ☐ Basic
- ☐ Advanced

## Member Services

- ☐ Join IEEE
- ☐ Establish IEEE Web Account

 Print Format[SEARCH RESULTS](#) [\[PDF Full-Text \(244 KB\)\]](#) [PREVIOUS](#) [NEXT](#)[DOWNLOAD CITATION](#)

VORTEX: Video retrieval and tracking from compressed multimedia databases-visual search engine

- Schonfeld, D. Lelescu, D.

Editor(s): Sprague, R.H., Jr.

Dept. of Electr. Eng. & Comput. Sci., Illinois Univ., Chicago, IL, USA

*This paper appears in:* Systems Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on

On page(s): 10 pp.

5-8 Jan. 1999

Maui, HI, USA

1999

ISBN: 0-7695-0001-3

Number of Pages: liii+341

References Cited: 14

INSPEC Accession Number: 6182119

**Abstract:**

Multimedia data is generally stored in compressed form in order to efficiently utilize the available storage facilities. Access to archives is dependent on our ability to browse compressed multimedia information-retrieval and tracking from coded video databases. In this paper, a novel visual search engine for video retrieval and tracking from compressed multimedia databases is proposed. The goal of the project is the implementation of a visual browser that operates in a distributed environment where users initiate video searches and retrieve relevant video information simultaneously from multiple video archives. Being presented with a query in the form of template images of objects, the system operates on the compressed video to find the images or video sequences where those objects are present and their positions in the image. Upon user's request, the system will decompress and display only the video sequences of interest.

**Index Terms:**

[video databases](#) [multimedia databases](#) [data compression](#) [video coding](#) [image retrieval](#) [search engines](#) [VORTEX](#) [video retrieval and tracking](#) [compressed multimedia](#) [visual search engine](#) [multimedia databases](#) [coded video databases](#) [video retrieval](#) [visual browser](#)

**Documents that cite this document**

Select link to view other documents in the database that cite this one.

[SEARCH RESULTS](#) [\[PDF Full-Text \(244 KB\)\]](#) [PREVIOUS](#) [NEXT](#)[DOWNLOAD CITATION](#)

[Advanced Search](#)

[Join IEEE](#) | [Web Account](#) | [New this week](#) | [OPAC Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Email Alerting](#)

[No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to Top](#)

Copyright © 2002 IEEE — All rights reserved

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

**Tables of Contents**

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

**Search**

- ☐ By Author
- ☐ Basic
- ☐ Advanced

**Member Services**

- ☐ Join IEEE
- ☐ Establish IEEE Web Account

 [Print Format](#)
[SEARCH RESULTS](#) [[PDF Full-Text \(64 KB\)](#)] [PREVIOUS](#) [NEXT](#)
**DOWNLOAD CITATION**

Word-based compression methods for large text documents

 - [Dvorsky, J.](#) [Pokorny, J.](#) [Snsel, V.](#)

Editor(s): Storer, J.A., Cohn, M.

Dept. of Comput. Sci., Olomouc Palacky Univ., Czechoslovakia

*This paper appears in:* Data Compression Conference, 1999.

Proceedings. DCC '99

On page(s): 523

29-31 March 1999

Snowbird, UT, USA

1999

ISBN: 0-7695-0096-X

IEEE Catalog Number: PR00096

Number of Pages: xv+566

References Cited: 3

INSPEC Accession Number: 6314316

**Abstract:**

Summary form only given. We present a new compression method, called WLZW, which is a word-based modification of classic LZW. The algorithm is two-phase, it uses only one table for words and non-words (so called tokens), and a single data structure for the lexicon is usable as a text index. The length of words and non-words is restricted. This feature improves the compress ratio achieved. Tokens of unlimited length alternate, when they are read from the input stream. Because of restricted length of tokens alternating of tokens is corrupted, because some tokens are divided into several parts of same type. To save alternating of tokens two special tokens are created. They are empty word and empty non-word. They contain no character. Empty word is inserted between two non-words and empty non-word between two words. Alternating of tokens is saved for all sequences of tokens. The alternating of tokens is an important piece of information. With this knowledge the kind of the next token can be predicted. One selected (so-called victim) non-word can be deleted from input stream. An algorithm to search the victim is also presented. In the decompression phase, a deleted victim is recognized as an error in alternating of words and non-words in sequence. The algorithm was tested on many texts in different formats (ASCII, RTF). The Canterbury corpus, a large set, was used as a standard for publication results. The compression ratio achieved is fairly good, on average 25%-22%. Decompression is very fast. Moreover, the algorithm enables evaluation of database queries in given text. This supports the idea of leaving data in the compressed state as long as possible, and to decompress it when it is necessary.

**Index Terms:**

[data compression](#) [text analysis](#) [query processing](#) [data structures](#) [sequences](#) [word-based compression](#) [large text documents](#) [WLZW](#) [two-phase algorithm](#) [data structure](#) [lexicon](#) [text index](#) [token length](#) [empty word](#) [empty non-word](#) [sequences](#) [search algorithm](#) [ASCII](#) [RTF](#)

Canterbury corpus database queries

---

**Documents that cite this document**

Select link to view other documents in the database that cite this one.

---

SEARCH RESULTS   [PDF Full-Text (64 KB)]   PREVIOUS   NEXT  
DOWNLOAD CITATION

---

Home | Log-out | Journals | Conference Proceedings | Standards | Search by Author | Basic Search |  
Advanced Search  
Join IEEE | Web Account | New this week | OPAC Linking Information | Your Feedback | Technical  
Support | Email Alerting  
No Robots Please | Release Notes | IEEE Online Publications | Help | FAQ | Terms | Back to Top

Copyright © 2002 IEEE — All rights reserved